Project Title

The Tuition Inquisition

Group Members

Alec Hubel, Yu Suzuki, David Eicke

Project Goals:

We set out to create a narrative-form tool to help our users to grasp the severity and explore the possible causes of rapidly increasing costs of higher education in the United States. We wanted to subvert the bombastic headlines on the topic and show that the problem is complex, requiring more than a simple, culpability-hurling explanation as often found in the media.

Related Work:

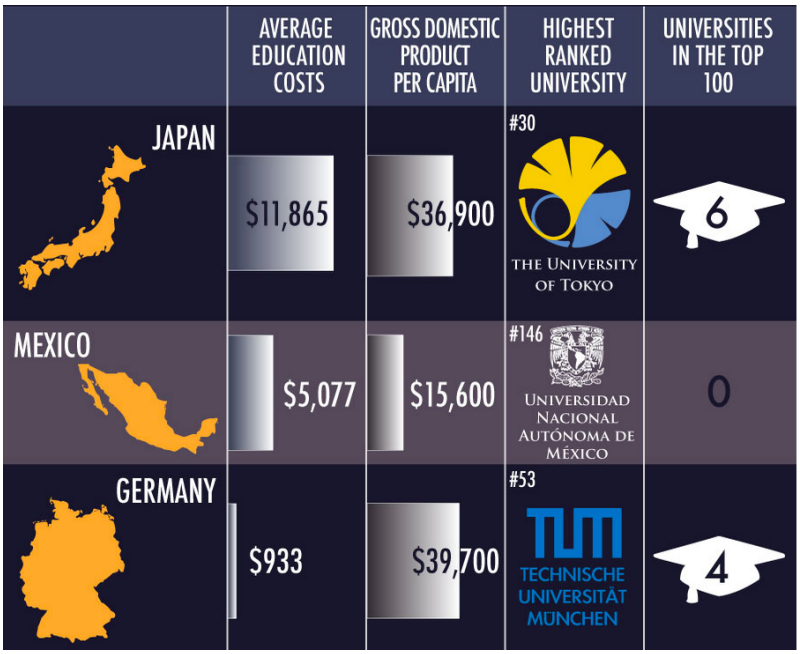
Excellence v equity

http://www.economist.com/news/special-report/21646985-american-model-higher-education-spreading-it-good-producing-excellence

The above article from *The Economist* planted the original seed for this project. In order to outline the differences between the American and European models of high education, the articles explains from a foundational supply-and-demand perspective why American education is so expensive. After reading this article, we came to the decision to build a visualization that helped to deconstruct the issues around college tuition.

College Tuition Around The World

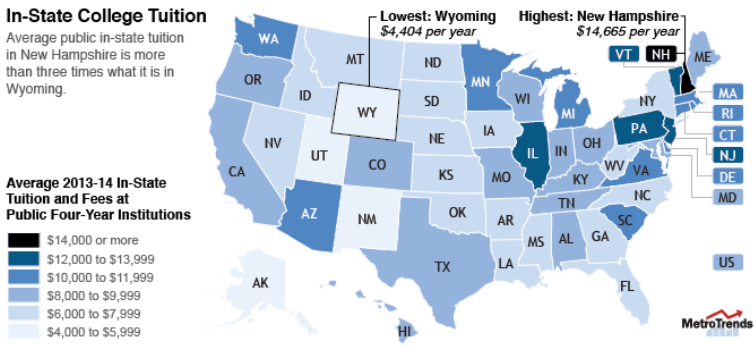
http://www.top10onlinecolleges.org/college-tuition/



This is an infographic comparing educational costs from several other nations. Countries are measured based on 4 metrics - average higher education costs, GDP per capita, the highest ranked university within that country, and the number of schools within that country that are ranked among the global Top 100 universities. From this visualization, we extracted data for our choropleth.

What the average college price really means for students

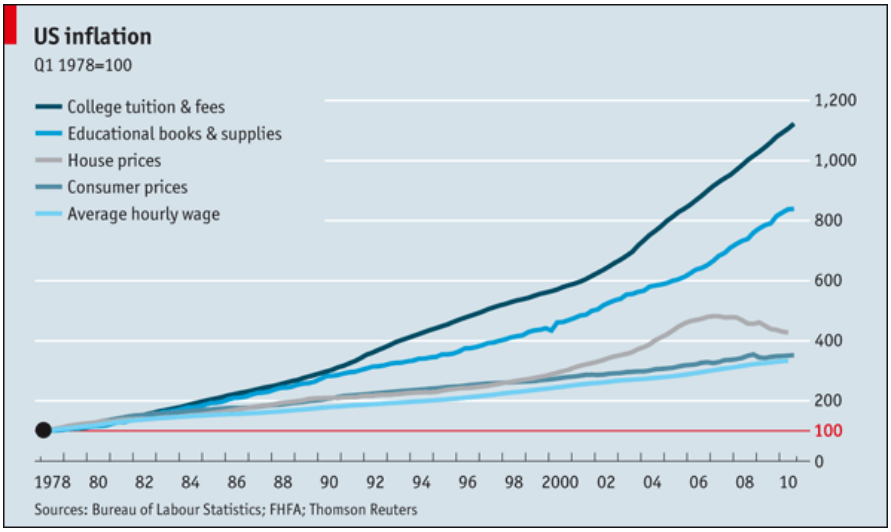
http://www.urban.org/urban-wire/what-average-college-price-really-means-students



The above visualization is taken from a report written by The Urban Institute, a social and economic policy think tank. In order to help show how college tuitions play out on a regional level, the report uses a choropleth to visualize state-by-state tuition prices. From this report, we took the inspiration to build our own choropleth, but focused on international tuition trends.

Higher education

http://www.economist.com/node/16960438?story_id=16960438



The above chart is taken from an article that was featured in *The Economist* that sought to put rising college tuitions in the context of other everyday costs. Although all day-to-day expenses are affected by inflation, we can see that educational costs are skyrocketing. This chart served as the inspiration for our graphic that compared the increase in college tuitions to easily recognizable goods (eggs, movie tickets, cars).

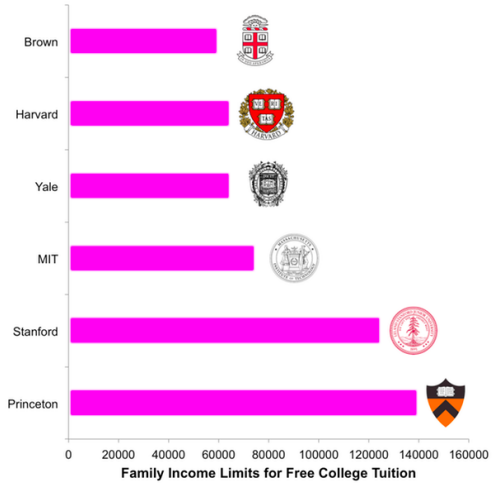
The Real Reason College Tuition Costs So Much

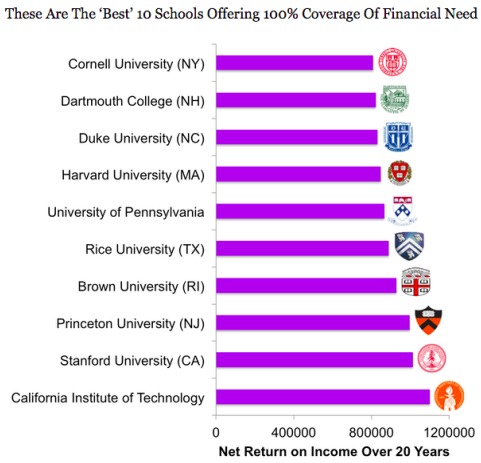
http://www.nytimes.com/2015/04/05/opinion/sunday/the-real-reason-college-tuition-costs-so-much.html

The above piece is an op-ed from the *New York Times* that attempts to uncover some of the drivers of tuition costs. After first trying to debunk the narrative that places responsibility on lower government funding, the author pins blame on increased administrative headcount. From this article, we took inspiration as well as data for our isotype graphic.

MIT, Harvard and Yale Among Colleges Offering Free Tuition

http://samuelwbennett.com/the-struggle/these-colleges-offer-free.html

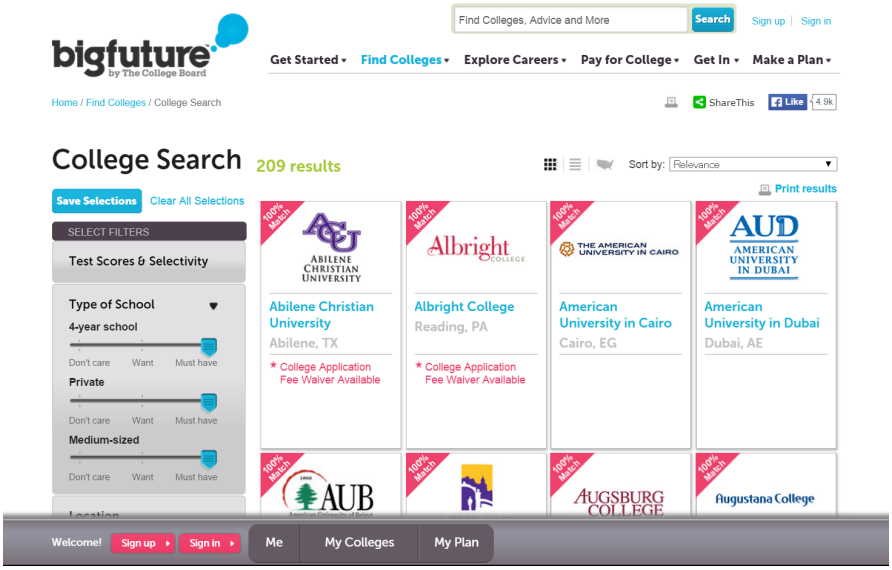




The above visualizations are from Samuel Bennett's blog, "get fast." In this particular blog post, he approaches the cost-of-college problem from the perspectives of of financial aid, and the potential return on investment from different schools. What we were most influenced by from these visualizations was the decision to use the school logos to help identify data points. This was eventually built in to our final dashboard.

Big Future

https://bigfuture.collegeboard.org/college-search



The College Board has rolled out a platform called "Big Future" that is targeted towards high school students who are searching for the right college for them. During our initial design process, we considered building a structure like this one for our exploratory dashboard.

However, we eventually came to the conclusion that a series of bars and sliders would not be the best way for us to present our data. Instead, we wanted to focus more on providing granular information on each school within our dataset.

Human groups of the earth
Chart from Die bunte Welt, image obtained from the class slide.

The above isotype visualization, taken from a class side (in turn taken from *Die bunte Welt*) is meant to visualize relative sizes of different ethnic groups. It uses figures of similar size and shape, differentiated only by color and outfit, to signify population sizes. We used the lessons learned from this image to inform the design of our own isotype visualization. More detail on our final design can be found below.

Exploratory Vistas - Ways to Become Acquainted with a Data Set for the First Time
Stephen Few

Exploratory Vistas

Ways to Become Acquainted with a Data Set for the First Time

Stephen Few, Perceptual Edge
Visual Business Intelligence Newsletter
July/August/September 2011

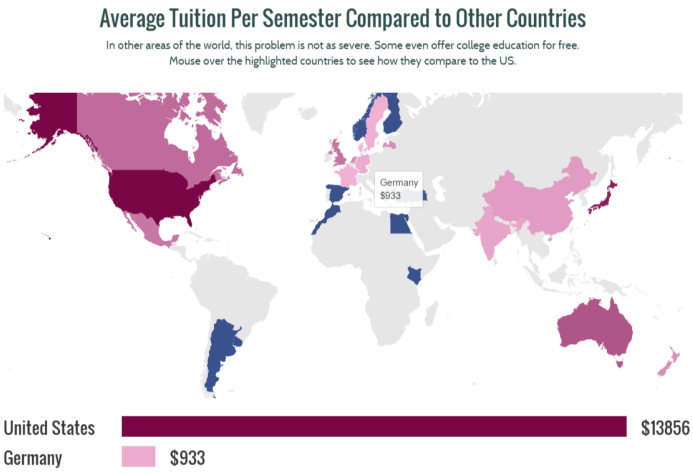
When you lay eyes on an unfamiliar territory for the first time, it always makes sense to get an overview before venturing into the thick of it. An unfamiliar data set is like an unknown land. Unless you have unlimited time to wander (who does?) and don't mind getting lost, it helps to study a map before starting out. If no map exists, then you should head for high ground to get the lay of the land.

This article from Stephen Few gives some useful insights for constructing an exploratory dashboard - which was the goal for the final portion of our project. Two insights in particular stood out for us. First, Few explains the importance of digesting contextual information (e.g. historical data, overarching trends, etc.) before diving in to data exploration. Because of this, we decided to have the exploratory portion of our project at the end of the page, so users would be forced to go through the smaller vignettes and learn about the problem space before dissecting the data. Second, Few emphasizes the importance of having multiple modalities coexist within a single dashboard. To follow this precept, we incorporated bar charts, iconography (via the school logos), and text-based information into the visualization.

Visualization Description:

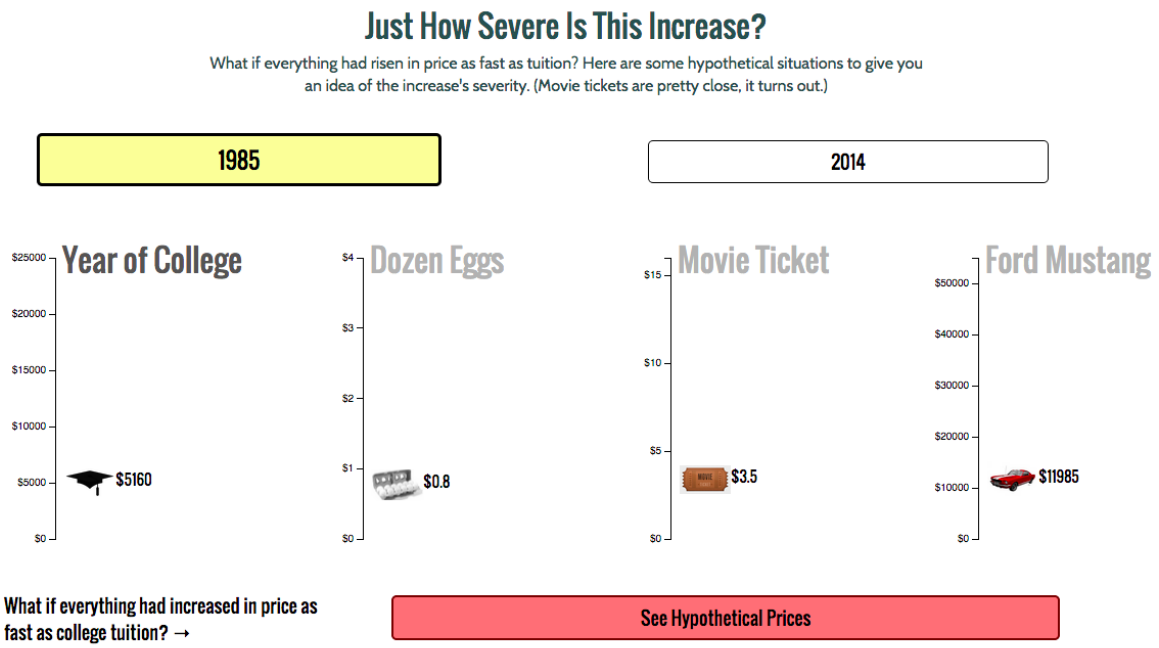
The visualization is a series of vignettes, each with a slightly different but related purpose. The introductory page is minimalist and simply introduces the problem space and how the visualization will approach it.

The first vignette places the United States within the context of the rest of the world. It shows that our average tuition dwarfs that of other nations, developed and developing. This was inspired by Yu, who, having lived all over the world, was genuinely curious about the US's high tertiary education costs. The color scheme for this choropleth uses a pink gradient, with darker hues representing more expensive countries and lighter hues representing cheaper countries. If a country offered higher education for free, we moved it outside of the pink gradient and colored it blue. We hoped that this would leverage preattentive visual properties, and draw immediate attention to the these countries.

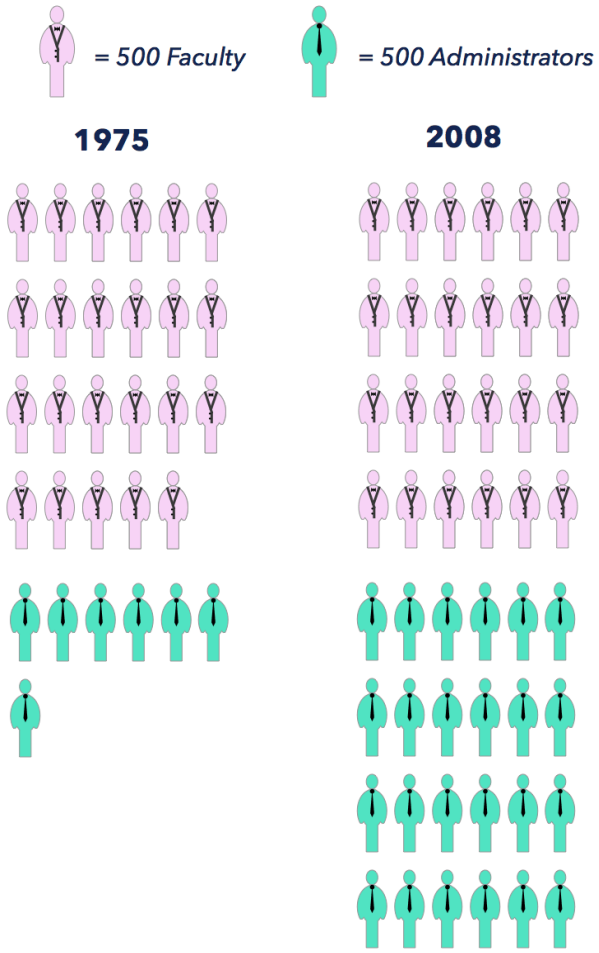


The second vignette illustrates the severity of the problem in a relatable way, using a hypothetical situation concerning three commonly known products that haven't changed much in nature and relative quality since 1985. We wanted to show what the prices of those products would be today, had their prices increased at the rate college tuition has increased. We got this from an article in the *New York Times* that asserted that a new car would cost $80,000 if its price had increased as quickly. We found this to be a number that didn't reflect reality very well, so we used the calculated value from our dataset.

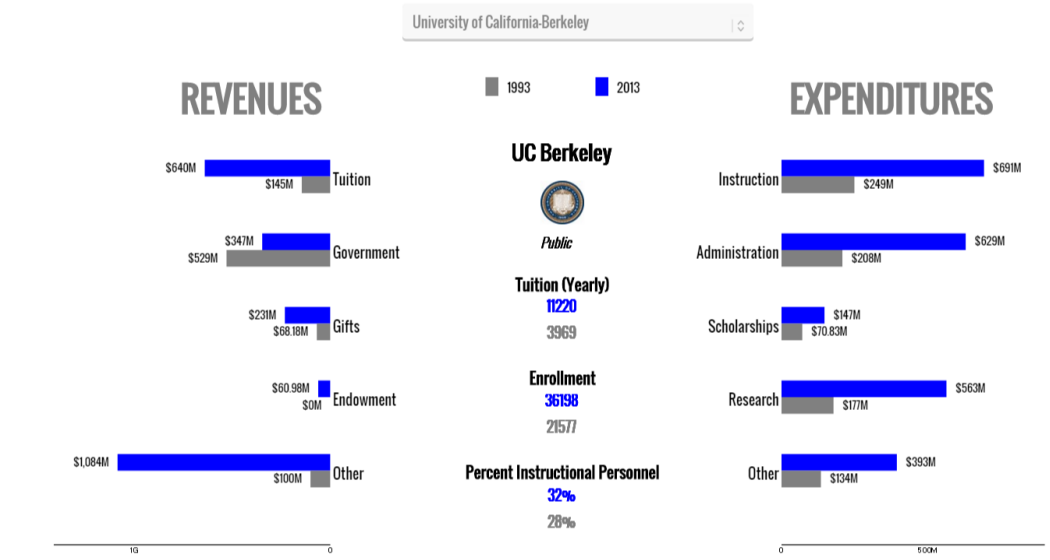
For this piece of our project, we tried to use several visual cues to help present our message. The two buttons at the top change formatting depending on which one is active - 1985 or 2014. The active button is given a background color and thicker border to highlight, while the inactive button has a thin border with no background coloring. The button on the bottom that is used to move the prices to their hypothetical state, is larger and shaded red, to signify its separation from the other two buttons. Additionally, this button is activated by a click, as opposed to a hover, in order to add a sense of intentionality to the experience. The text for the "Year of College" chart is made darker than the three other charts to signify it's relative importance. As the icons representing the four goods move from 1985 to 2014, and then from 2014 to the hypothetical values, a faded version of the icon is left behind the previous states. This is so that the user can make quick comparisons between the three values.

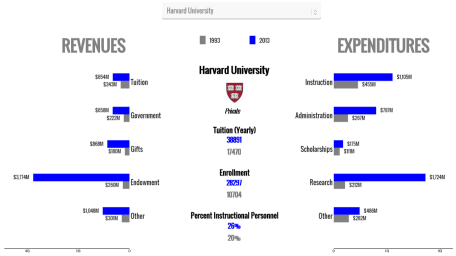


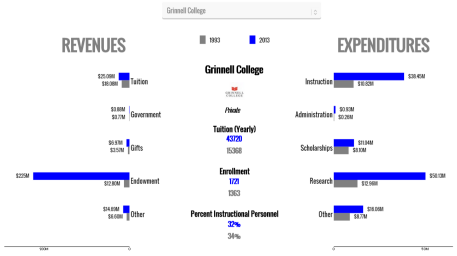
The third vignette demonstrates a possible factor in rising cost, also gleaned from the *New York Times* article on the topic. One professor did a study of the CSU system and found that while the number of instructional faculty had increased by only 4%, the number of administrative faculty had increased 221%. We decided to illustrate this increase with isotypes to get the biggest visual impact and retain the "personality types" of the actors involved. To differentiate the faculty and administrator isotypes, we added some small, yet distinct wardrobe details. The faculty were adorned with bow ties and sweater vests, while the administrators are wearing straight ties. Because this distinction would likely not be enough, we also shaded the faculty isotypes pink, while the administrator isotypes are shaded green.

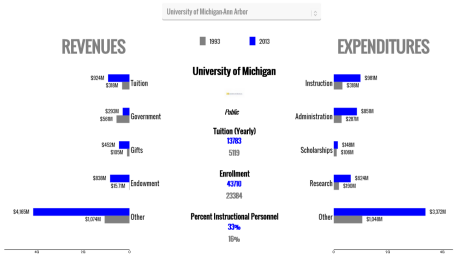


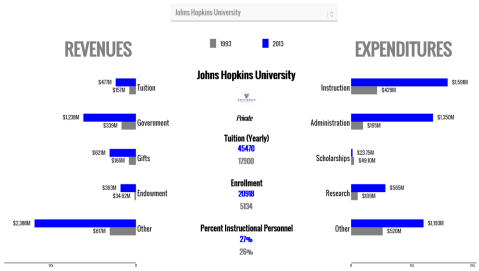
Our fourth vignette was our main dashboard and the “exploratory” part of our visualization. (The prior vignettes had been more about “presentation.”) We decided to mimic a flow-in, flow-out model of revenue streams and expenditures. We thought the best way to do this would be to have one set of horizontal bar graphs on the left (revenues) and one set on the right (expenditures), with some information about the college in between. Everything having to do with current data (2013) was in blue, while every data point having to do with the past (1993) was in gray. The user could select the college he or she was interested in from a dropdown menu. Once selected, the revenue and expenditure data would update accordingly.











In the above screen shots, we have extracted some representative cases from our data set. The visualization for UC Berkeley (top center) and the University of Michigan (bottom left) highlight some of the trends for large public universities. These schools tend to have dropping government funding (from 1993 to 2013) and particularly large “other” bars on both the revenue and expenditure sides. This was likely because that these schools dedicate substantial finances to sports, licensing, real estate, and other non-academic pursuits. Harvard University and Grinnell College (middle left and right) are notable because of their oversized endowments.

Data:

Data for the initial choropleth was taken from a report produced by Higher Education Strategy Associates (a higher education think tank) entitled *Global High Education Rankings 2010*, that outlines overarching trends in international higher education. A visualization based on this report is outlined in the "related work" section. We would have liked to dig up more data for this section (our map has a lot of gray on it) but were unable to find data sets that provided this information.

Data for the isotype section was taken from an analysis that was performed by a professor at California Polytechnic University, Pomona. This study was noted in the aforementioned *New York Times* piece, "The Real Reason College Tuition Costs So Much."

Data for the final dashboard was collected from the Integrated Postsecondary Education Data System (IPEDS), a database maintained by the Department of Education that holds data for every school in the United States that is eligible for Title IV funding. Data is available for as far back as 1980, and as recently as 2013.

Data for the comparison graphic was taken from a variety of sources. The cost of college was averaged from data collected via IPEDS. Prices for the Ford Mustang was taken from a website, howstuffworks.com, that listed historical information for this car model. Historical egg prices was collected from the bureau of labor statistics. Movie ticket prices were taken from the website for the National Association of Theatre Owners.

Tools

To download and properly format the data, we used a combination of python scripts and excel work. We performed some initial data exploration and analysis in Tableau. The final visualization was constructed using a combination of HTML, CSS, & JavaScript (jQuery and D3 libraries).

Steps Performed: The steps we used to construct our visualization are as follows:

1. Access and combine relevant data
2. Decide on a manageable number of schools
3. Explore the data ourselves for trends
4. Develop some visualization-based narratives to accompany our exploratory narrative
5. Come up with best ways to facilitate exploration of data
6. Find related works
7. Wireframe the interface
8. Coding
9. User Testing on fellow students
10. Reformatting some of the visualizations based on feedback

Results:

While we were overall pretty happy with the result, we did find several issues of concern in user testing, and there was one feature we didn't get to that we would have liked to include.

Starting from the beginning, the choropleth was fairly simple for people to explore. However, vast sections of the map were without data. We would like to find more, but we also would want it to be reliable. On top of this, there was one user who asked about the normalization of the cost by purchase power in the individual country. While we were consistent in using USD throughout, the buying power of the dollar is different in other countries, and it would have been smart of us to adjust the data accordingly, or at least make it a toggle option.

The comparison visualization did not have the impact we wanted it to have, mostly because we feel that the prices of the products we chose to use were not very shocking when compared to their hypotheticals. (For eggs, in particular, the regional variation and quality variation inflate or deflate the price more than 20 years' worth of time.) The UI was also difficult for people, who seemed to gloss over the fact that there was a button to see the hypothetical prices. Instructions to mouse over the years were present but users rarely read them. Users also attempted to manipulate the actual icons, believing that they could drag them. Originally, we had "One Year of College" on the top by itself, but we reconsidered that and moved it down to be next to the other items that it was being compared to. After the showcase, however, we realized that we might not even need tuition in there at all. Perhaps we could just have had a button to say "Click to see prices if they'd increased as fast as tuition.

The isotype ended up being the easiest portion of the visualization to understand. We got a lot of positive feedback from this part. We considered reflecting the imbalance of income between your average administrator and your average faculty by making the faculty isotypes thinner. However, we couldn't find data to back this up. We also thought it might interfere with the visual impact of the isotype with more space in between one half of the graphics (the alternative being that the space between them is the same and one set is narrower than the other). One interesting piece of feedback we got was that a user was perceiving this as an accusation of bad practices rather than an unbiased count. We wanted to present this as possibly part of an explanation of rising costs, rather than an editorial saying "they shouldn't be hiring all these administrators!"

The dashboard feature was the most interesting part for many of our users, as we had hoped and anticipated. Unfortunately, we did not get the time to create our other feature, which would be a toggled functionality where users get to compare two selected schools with one another. A couple people asked for this feature, so our anticipation of this desire seems to be correct. Problems with the dashboard included:

1. *Tendency to ignore data in the center of the visualization* - Because there was a lot of text in one area (labels for the bars, school names, numbers, subheadings), we believe the information got lost. A solution for the future might be to take that information out of the center and put it in a box by itself at the top or on the side.
2. *Association between horizontally aligned items* - We had anticipated that this might be a problem, as sets of bars lined up across from each other tend, not only to be related, but to have the same axis. We had offset the graphs slightly in an earlier version, but it just looked like a mistake. In future iterations, we might try different colored backgrounds or possibly different shades to communicate that the bars across from each other are not directly related to each other. We also failed to communicate the 'flow in, flow out' model. It appeared that the bars were all radiating "out" of the center. We might fix this by moving the axis to the left for both revenue and expenditure. That way, the "flow" of the visualization will be moving in the same direction.
3. *Too much choice* - When users sat down with the dashboard, most immediately went to investigate their alma mater. From there, they didn't really know where to go, so they

visited famous schools. If we had pulled out some interesting ones to begin with, people would get the point faster and have some direction after they've satisfied their curiosity.

Link to Visualization

http://people.ischool.berkeley.edu/~dreicke/InfoVizTuition/Framework/

Team Member Contribution:

(Everyone contributed to concept and overall structure, as well as the writing of this report.)

DAVID EICKE	**ALEC HUBEL**	**YU SUZUKI**
Most of the D3 for dashboard	Main IPEDS and Forbes data scraping	World Map visualization
CSS styling	All data sorting and wrangling	International Tuition data research
Data research for 1985 prices	Common-goods comparison visualization	Drop-down menu code for dashboard
Isotypes graphic	Coding center data for dashboard	Function for converting numbers like $232423000 into $232M for readability
Most written copy and instructions	Color scheme	Color Key for dashboard
		iFrames structure

Thumbnail

See attached

Github Repository

https://github.com/ahubel/InfoVizTuition